Distinguishability, classical information of quantum operations

Dong Yang*
Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China
(Dated: February 1, 2008)

A basic property of distinguishability is that it is non-increasing under further quantum operations. Following this, we generalize two measures of distinguishability of pure states—fidelity and von Neumann entropy, to mixed states as self-consistent measures. Then we extend these two measures to quantum operations. The information-theoretic point of the generalized Holevo quantity of an ensemble of quantum operations is constructed. Preferably it is an additive measure. The exact formula for $SU(2)$ ensemble is presented. With the aid of the formula, we show Jozsa-Schlienz paradox that states as a whole are less distinguishable while all pairwise are more distinguishable in an ensemble of quantum states, also occurs in an ensemble of quantum operations, even in the minimal dimensional case $SU(2)$ ensemble.

PACS numbers: 03.67.-a, 03.65.Ta

I. INTRODUCTION

Quantum nonorthogonality is one of the basic features of quantum mechanics. Unlike the distinct states in classical physics, it is impossible to discriminate perfectly between nonorthogonal states if only one copy is provided. Motivated by this simple fact, there have been attempts to define measures to quantify the degree of distinguishability between quantum states. A well-known measure is the fidelity between two quantum states [1] that describes: the more similar, the less distinguishable they are. Recently, von Neumann entropy is suggested to measure the distinguishability for an ensemble of pure states from the information-theoretic point [2]: the more classical information they can communicate, the more distinguishable they are. Indeed, von Neumann entropy is the classical information capacity communicated by an ensemble of pure states [3]. Quantum states and quantum dynamics are two parts of quantum mechanics. Both quantum states and dynamics are fundamental physical resources [4]. A general quantum dynamical process is described by a quantum operation, a completely positive trace-preserving linear map (CPT) [5, 6]. Quantum nonorthogonality occurs not only in quantum states but also in quantum operations. It is possible that two operations can not be discriminated exactly if only one operation is performed. Distinguishability between two unitary operations is discussed in [7, 8, 9]. However, little is known about the distinguishability of general quantum operations. Motivated by this, we try to quantify the distinguishability of quantum operations.

How can one introduce a measure of distinguishability for quantum operations? A natural approach exists if two key points are noticed. One is the basic property of distinguishability that it is non-increasing under further quantum operations and the other is that distinguishability of quantum operations can be defined based on distinguishability of quantum states. More specifically speaking, distinguishability of pure states is extended to that of mixed states, and then distinguishability of quantum operations is proposed as the largest distinguishability of the output states (mixed states). In this paper, we discuss two measures of distinguishability: one is fidelity and the other is Holevo quantity [10]. The rest of the paper is organized as follows. Section II recalls the notions of CPT map and properties of Holevo quantity that is useful to study distinguishability. Section III first reviews two measures of distinguishability of pure states—fidelity and von Neumann entropy, then generalizes to those of mixed states as self-consistent measures. Section IV explores the corresponding measures for distinguishability of quantum operations. Several propositions are proved and the information-theoretic explanation of Holevo quantity is constructed. Further the formula to calculating the distinguishability of $SU(2)$ ensemble is provided. Employing the formula, a counter-intuitive phenomenon is discovered analogous to Jozsa-Schlienz paradox in [2]. Section V concludes with summary.

*Electronic address: dyang@ustc.edu.cn

II. CPT MAP AND HOLEVO QUANTITY

Before discussing distinguishability, we briefly review the CPT map and Holevo quantity that are used in the latter sections.

A. CPT Map

A general quantum states is described by a positive Hermite operator with trace one. A general quantum dynamical process is described by a quantum operation, a completely positive trace-preserving linear map (CPT)[5, 6], $\rho_{out} = \mathcal{E}(\rho_{in})$. 'Linear' means $\mathcal{E}(p_1\rho_1 + p_2\rho_2) = p_1\mathcal{E}(\rho_1) + p_2\mathcal{E}(\rho_2)$. 'Trace-preserving' means $tr\mathcal{E}(\rho) = tr\rho$. 'Positive' means $\mathcal{E}(\rho) \geq 0$ if $\rho \geq 0$. 'Completely positive' means $\mathcal{E} \otimes I$ is positive where I is the identity map on any dimensions. There are two useful representations of a CPT map $\mathcal{E}$. One is the operator-sum representation, $\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger$, where A_i are operators satisfying $\sum_i A_i^\dagger A_i = I$. The other is unitary representation: a CPT map can be written as a unitary evolution on an extended system AB followed by a partial trace over B, $\mathcal{E}^A(\rho^A) = tr_B U^{AB} \rho^A \otimes (|0\rangle\langle 0|)^B U^{AB\dagger}$. For more details and the relation between the two representations, the reader is refereed to [5, 6].

B. von Neumann Entropy and Holevo quantity

Von Neumann entropy of a density operator is defined as $S(\rho) = -tr\rho \log \rho$, where logarithm is to base 2. It has a remarkable property known as strong subadditivity inequality [11], $S(\rho_{12}) + S(\rho_{23}) \geq S(\rho_{123}) + S(\rho_2)$, with any tripartite state ρ_{123} on Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$.

An ensemble of mixed states is denoted as $\{\rho_i, p_i\}$ where ρ_i are quantum states and $p = \{p_i\}$ is a probability distribution satisfying $p_i \geq 0, \sum_i p_i = 1$. Holevo quantity for an ensemble of mixed states is defined as,

$$\chi(\{\rho_i, p_i\}) = S(\sum p_i \rho_i) - \sum p_i S(\rho_i). \tag{1}$$

For an ensemble of pure states, Holevo quantity is reduce to von Neumann entropy. Holevo quantity has two properties [12, 13, 14].

Proposition: Holevo quantity is non-increasing under trace operation and under quantum operation,

$$\chi(\{\rho_i^{AB}, p_i\}) \geq \chi(\{\rho_i^A, p_i\}), \tag{2a}$$
$$\chi(\{\rho_i, p_i\}) \geq \chi(\{\mathcal{E}(\rho_i), p_i\}), \tag{2b}$$

where $\rho_i^A = tr_B \rho_i^{AB}$ and $\mathcal{E}$ is a CPT map.

The proof comes from the strong subadditivity inequality of von Neumann entropy [12, 13, 14].

III. DISTINGUISHABILITY OF QUANTUM STATES

Now we investigate the distinguishability measure of quantum states. Intuitively, the word *distinguishability* means to what extent things are distinct from each other. As a meaningful measure of distinguishability among the objects, the quantity cannot be increased by the same further processing, otherwise we are in dilemma that the quantity tends to infinity that is clearly meaningless. In quantum mechanics, the further processing is quantum operation described by CPM. So we argue that a reasonable measure of distinguishability defined in quantum mechanics should satisfy the constraint: *a measure of distinguishability is non-increasing under any further quantum operation.* We will see how this constraint leads us to define a self-consistent measure of distinguishability.

A. Fidelity

From the intuition that the more alike the less distinguishable, distinguishability of two states is measured by fidelity. First it is defined on two pure states, then is extended to mixed ones under the constraint.

Suppose two pure nonorthogonal states $|\phi\rangle$ and $|\psi\rangle$, the overlap or fidelity is defined as $F(\phi, \psi) = |\langle\phi|\psi\rangle|$ and measures to what extent the two states behave alike. Indeed, $|\langle\phi|\psi\rangle|^2$ is the probability that $|\phi\rangle$ passes the test of 'being state $|\psi\rangle$'. Note that the fidelity is invariant under unitary evolution which means that the two states behave the

same alike as the beginning if they evolve under the same unitary dynamics. The more alike, the less distinguishable they are. So fidelity can be used to mark distinguishability. However, before that we should check whether fidelity is non-decreasing under quantum operation as distinguishability is non-increasing under quantum operation. Under general quantum dynamics, pure states evolve into mixed states. As a natural requirement, it is necessary to define fidelity between two mixed states such that it could be reduced to the pure case when they are pure. Note that mixed states can be purified to entangled pure states (purification) with the aid of an auxiliary subsystem, and can be regarded as partially tracing the auxiliary subsystem over the purification. It is possible to define fidelity of two mixed states as that of purifications. Given a mixed state, purification is not unique—infinite purifications exist. How do we select a pair of purifications for definition? The answer also comes from the constraint that distinguishability is non-increasing under quantum operations. As partial-tracing is also a quantum operation, only one way is left to define fidelity of mixed states as,

$$F(\rho_1, \rho_2) = \max |\langle \Phi_1 | \Phi_2 \rangle|, \tag{3}$$

where the maximum is taken over all purifications $|\Phi_1\rangle, |\Phi_2\rangle$. The physical meaning is that $F(\rho_1, \rho_2)$ is the worst distinguishability between purifications. This definition is also accordance with the meaning of distinguishability: the more we learned, the better we can distinguish. But does $F(\rho_1, \rho_2)$ satisfy the requirement of distinguishability? Indeed, the definition of (3) is self-consistent, which means that $F(\rho_1, \rho_2)$ is non-decreasing under further quantum operation. Indeed, Jozsa proved the following proposition[1].

Proposition: $F(\rho_1, \rho_2)$ is non-decreasing under tracing subsystem operation and under quantum operation,

$$F(\rho_1^A, \rho_2^A) \geq F(\rho_1^{AB}, \rho_2^{AB}), \tag{4a}$$

$$F(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)) \geq F(\rho_1, \rho_2), \tag{4b}$$

where $\rho_i^A = tr_B \rho_i^{AB}$.

It is further showed [1] that$F(\rho_1, \rho_2) = tr[(\sqrt{\rho_1}\rho_2\sqrt{\rho_1})^{1/2}]$,where $\{tr[(\sqrt{\rho_1}\rho_2\sqrt{\rho_1})^{1/2}]\}^2$ is introduced as 'transition probability' for mixed states by Uhlmann in [15]. For more details about $F(\rho_1, \rho_2)$, the readers are refereed to [1].

B. von Neumann Entropy and Holevo quantity

The distinguishability measure by fidelity is defined for two states. How can we define the measure for an ensemble of states? From the information-theoretic point that the more distinguishable of a set of states are, the more information they can communicate, Jozsa and Schlienz [2] proposed that von Neumann entropy can be used to quantify distinguishability of an ensemble of pure states, where the distinguishability measure of $E = \{|\phi_i\rangle, p_i\}$ is defined as $D(E) = S(\sum p_i |\phi_i\rangle\langle\phi_i|)$. Indeed, von Neumann entropy can be explained as the classical information capacity communicated by the ensemble of pure states [3]. As a self-consistent definition, a distinguishability measure should be defined on mixed states because the measure should be non-increasing under further quantum operation and pure states generally evolve into mixed ones. An explicit generalization of von Neumann entropy is Holevo quantity in the mixed case $E = \{\rho_i, p_i\}$,

$$D(E) = \chi(\{\rho_i, p_i\}). \tag{5}$$

The Holevo quantity is non-increasing under further quantum operation. More importantly, it does measure the classical information capacity for an ensemble of mixed states [16]. So the Holevo quantity is indeed a self-consistent measure based on the information-theoretic point.

Remarks: (1) Intuitively, the more distinguishable each pair of quantum states, the larger distinguishability of the ensemble. However, a counter-intuitive fact known as Jozsa-Schlienz paradox is found in [2], where the states becomes all pairwise more distinguishable while the distinguishability of the ensemble measured by von Neumann entropy decreases. This phenomenon shows that distinguishability measured by von Neumann entropy is a global property of an ensemble of pure states and not an accumulative local property of pairs of constituent states [2]. (2) From the information-theoretic explanation, von Neumann entropy is a quantity in the asymptotic meaning [3, 6], in which physical quantities is generally defined by the regularization form, i. e. $\lim_{n\to\infty} D(E^{\otimes n})/n$. Just because von Neumann entropy and Holevo quantity are additive over tensor product of ensembles, the regularization is reduced to one copy. We are faced with the regularization problem when we define the distinguishability measure of operations.

C. Orders of Different Measures

Both fidelity and von Neumann entropy (Holevo quantity) can be used to measure the distinguishability of an ensemble with two states. For pure states, distinguishability measured by fidelity and von Neumann entropy give

the same order, i. e. if ensemble E_1 is more distinguishable than E_2 measured by fidelity, then it is also the case when measured by von Neumann entropy and vice versa. This can be easily seen that von Neumann entropy is monotonously dependent on the inner product of two states. However, for general mixed states it is possible that E_1 is more distinguishable than E_2 measured by fidelity while E_1 is less distinguishable than E_2 by Holevo quantity. A simple situation is: suppose two ensembles $E_1 = \{\rho_1, \rho_2, p_1, p_2\}$ and $E_2 = \{|\phi_1\rangle, |\phi_2\rangle, p_1, p_2\}$ where $|\phi_1\rangle$ and $|\phi_2\rangle$ are the optimal purifications to achieve $F(\rho_1, \rho_2) = |\langle\phi_1|\phi_2\rangle|$. From *Eq.*(2a), $\chi(E_1) \leq \chi(E_2)$ holds. For generic mixed states, $\chi(E_1)$ is strictly less than $\chi(E_2)$. As von Neumann entropy is continuously dependent on $\langle\phi_1|\phi_2\rangle$, $|\langle\phi_1|\phi_2\rangle|$ can be made a little larger by deforming $|\phi_1\rangle$ and $|\phi_2\rangle$ while the inequality $\chi(E_1) < \chi(E_2)$ still holds. So distinguishability measured by fidelity is not always inconsistent with that by Holevo quantity. The fact that different measures give different orders of distinguishability implies that a particular measure just reflects a particular property of distingushability in quantum states.

IV. DISTINGUISHABILITY OF QUANTUM OPERATIONS

Definition: An ensemble of quantum operations is defined as an ensemble $\{\mathcal{E}_i, p_i\}$, where $\mathcal{E}_i$ are CPT maps and $p = \{p_i\}$ is a probability distribution. When all $\mathcal{E}_i$ are unitary operators U_i, we call the ensemble $\{U_i, p_i\}$ unitary ensemble.

Suppose we are given one of black boxes that perform operations $\mathcal{E}_i$ with probability p_i and required to identify which operation the given box performs. Of course, if the black box can be inquired infinite times, that is to say the same operation can be performed infinitely, we can identify what the operation is by quantum operation tomography [18]. If we can inquire the box only once, to what extent can we tell which operation the box performs? But before that, what is the meaning of 'extent'? Therefore, it is necessary to define the measure of distinguishability— the degree of distinguishability between quantum operations. To distinguishing quantum operations, the only way is to input a state to the black box and the evolution is inferred by the output state. So distinguishability for an ensemble of quantum operations is reduced to distinguishability of the ensemble of output states. As there exist different ensembles of output states for different input states, distinguishability of the ensemble of quantum operations is defined as the maximal distinguishability of the ensemble of output states. The corresponding input state is the optimal one to distinguish the quantum operations as well as possible. Note also that given a CPT $\mathcal{E}_i$, we are allowed to perform any operation of the form $\mathcal{E}_i \otimes I$ where I is the identity operator acting on any dimensions. So it is possible that an entangled state including an auxiliary subsystem is better for distinguishing quantum operations. An explicit example is that $\{I, \sigma_x, \sigma_y, \sigma_z\}$ can be distinguished exactly by a maximally entangled state $|\phi\rangle = 1/\sqrt{2}(|00\rangle + |11\rangle)$, where $\sigma_x, \sigma_y, \sigma_z$ are Pauli operators. Here $\{\mathcal{E}_i, p_i\}$ is just a brief notation that actually means $\{\mathcal{E}_i \otimes I, p_i\}$.

In this section, parallel to distinguishability of quantum states measured by fidelity and Holevo quantity, the corresponding distinguishability of quantum operations is studies respectively.

A. Fidelity

In this subsection, we discuss fidelity of two quantum operations. In [8], distinguishability between two unitary operators U_1, U_2 is defined by the minimal fidelity, $F(U_1, U_2) = \min_{|\phi\rangle} F(U_1 \otimes I|\phi\rangle, U_2 \otimes I|\phi\rangle)$. From the same reason as quantum states, a self-consistent measure should be defined on general quantum operations.

Definition The fidelity of two quantum operations between $\mathcal{E}_1$ and $\mathcal{E}_2$ is defined as

$$F(\mathcal{E}_1, \mathcal{E}_2) = \min_{\rho} F(\mathcal{E}_1 \otimes I(\rho), \mathcal{E}_2 \otimes I(\rho)). \tag{6}$$

Before we prove that $F(\mathcal{E}_1, \mathcal{E}_2)$ is non-decreasing under further quantum operation, we show that minimization can be obtained over a pure state entangled with a finite-dimensional system.

Lemma 1 Given two quantum operations $\mathcal{E}_1$ and $\mathcal{E}_2$ acting on d-dimensional system A. $F(\mathcal{E}_1, \mathcal{E}_2)$ can be achieved by minimization over pure states entangled with a d-dimensional auxiliary system B at most. That is

$$F(\mathcal{E}_1, \mathcal{E}_2) = \min_{|\phi\rangle \in d \otimes d} F(\mathcal{E}_1 \otimes I_d(\phi), \mathcal{E}_2 \otimes I_d(\phi)) \tag{7}$$

Proof. For a general mixed state ρ^{AB} on $d \otimes n$-dimensional space, it can be always purified into a pure state $|\Phi\rangle^{ABC}$ with another auxiliary system C, where $|\Phi\rangle^{ABC}$ can be written in its Schmidt decomposition splitting between A and BC,

$$|\Phi\rangle^{ABC} = \sum_{i=0}^{d-1} \sqrt{\lambda_i} |i\rangle^A \otimes |e_i\rangle^{BC}, \tag{8}$$

in which $|e_i\rangle^{BC}$ are orthogonal states spanning d-dimensional space. From $Eq.$(4a), we can get

$$\begin{aligned} F(\mathcal{E}_1^A \otimes I^B(\rho^{AB}), \mathcal{E}_2^A \otimes I^B(\rho^{AB})) &= F(tr_C \mathcal{E}_1^A \otimes I^{BC}(\Phi^{ABC}), tr_C \mathcal{E}_2^A \otimes I^{BC}(\Phi^{ABC})) \\ &\geq F(\mathcal{E}_1^A \otimes I^{BC}(\Phi^{ABC}), \mathcal{E}_2^A \otimes I^{BC}(\Phi^{ABC})). \end{aligned} \tag{9}$$

Notice that Φ^{ABC} can be transform to the standard form $\sum_{i=0}^{d-1} \sqrt{\lambda_i}|i\rangle^A \otimes |i\rangle^{BC}$ under unitary operator $I^A \otimes U^{BC}$ without varying $F(\mathcal{E}_1^A \otimes I^{BC}(\Phi^{ABC}), \mathcal{E}_2^A \otimes I^{BC}(\Phi^{ABC}))$. So the minimization can be considered over pure states in the $d \otimes d$-dimensional Hilbert space.

Proposition 1 $F(\mathcal{E}_1, \mathcal{E}_2)$ is non-decreasing under further quantum operation,

$$F(\mathcal{N} \circ \mathcal{E}_1, \mathcal{N} \circ \mathcal{E}_2) \geq F(\mathcal{E}_1, \mathcal{E}_2), \tag{10}$$

where $\mathcal{N}$ is a general quantum operation.

Proof. The proof is immediately obtained from $Eq.$(4b),

$$F(\mathcal{N} \circ \mathcal{E}_1 \otimes I_d(\phi), \mathcal{N} \circ \mathcal{E}_2 \otimes I_d(\phi)) = F(\mathcal{N} \otimes I(\mathcal{E}_1 \otimes I_d(\phi)), \mathcal{N} \otimes I(\mathcal{E}_2 \otimes I_d(\phi)) \geq F(\mathcal{E}_1 \otimes I_d(\phi), \mathcal{E}_2 \otimes I_d(\phi)). \tag{11}$$

The problem of distinguishing quantum operations is different from that of quantum states though they seem much alike. In [8], a distinct property is demonstrated that for any two unitary operations U_1 and U_2 there always exists a finite number N such that $U_1^{\otimes N}$ and $U_2^{\otimes N}$ are perfectly distinguishable although they were not in the single-copy case. Recall that it is not the case for two nonorthogonal states. If two states are not discriminated perfectly with one copy, they are not with any finite copies. The input state entangled with different subsystems on which every U performs improves the distinguishability between $U_1^{\otimes N}$ and $U_2^{\otimes N}$. The result also implies that any number unitary operations can be exactly distinguished if enough finite copies are provided. An upper bound on copies is $N = \sum_{k=1}^{m-1} N^{(k)}$, where $N^{(k)}$ is the k-th number in the decreasing sequence of $\{N_{ij}, i \neq j\}$ and N_{ij} is the minimal copies that are required to distinguish U_i and U_j from the set $\{U_i, i = 1, \cdots, m\}$. The reason is that we can always use $N^{(k)}$ copies to exclude one possible and if the possible is excluded, the number of required copies does not appear in later exclusion. So we can exclude all possibilities arbitrarily. For example, let us denote the possible operation U_i. First, we use N_{12} copies to exclude the possibility of U_1 or U_2 by proper input state. If the outcome is in favor of U_1, the possibility of U_2 is excluded. Then we are left with $m-1$ possibilities and continue in this way until there is left only one possible. Here we don't optimize the problem. Indeed, much better upper bound can be obtained. Is it true that general operations can be perfectly distinguished with finite copies? The answer is negative. In the following, we give two instances.

A CPT map is called entanglement breaking if $\mathcal{E}^A \otimes I^B(\rho^{AB})$ is always separable for any ρ^{AB}. A CPT map is entanglement-breaking if and only if it can be written in the form [17]

$$\mathcal{E}(\rho) = \sum_i |\phi_i\rangle\langle\phi_i| tr(|\psi_i\rangle\langle\psi_i|\rho), \tag{12}$$

where $|\phi_i\rangle$ are normalized and $|\psi_i\rangle$ are not necessarily normalized but satisfy $\sum_i |\psi_i\rangle\langle\psi_i| = I$.

Ex 1: If two unitary operations $U_1, U_2 \in SU(d)$ are not perfectly distinguishable in the single-copy case, then two EB $\mathcal{E}_1 = U_1 \circ \mathcal{E}, \mathcal{E}_2 = U_2 \circ \mathcal{E}$ are not perfectly distinguishable in any finite-copy case, in which $\mathcal{E}$ is an EB map.

Proof.

$$\mathcal{E}_{1(2)}(\rho_{1(2)}) = \sum_i U_{1(2)}|\phi_i\rangle\langle\phi_i|U_{1(2)}^\dagger \otimes |tr(|\psi_i\rangle\langle\psi_i|\rho_{1(2)}), \tag{13}$$

Here $\rho_{1(2)}$ are bipartite states with ancilla system. Since $U_1|\phi_i\rangle$ is not orthogonal to $U_2|\phi_i\rangle$, if $\mathcal{E}_1(\rho_1) \perp \mathcal{E}_2(\rho_2)$, then $tr(|\psi_i\rangle\langle\psi_i|\rho_1) \perp tr(|\psi_i\rangle\langle\psi_i|\rho_2)$. As $\{|\psi_i\rangle\langle\psi_i|\}$ is a generalized measurement, it means that $\rho_1 \perp \rho_2$.

$$\begin{aligned} \rho_{1(2)} &= \mathcal{E}_{1(2)} \otimes \mathcal{E}_{1(2)} \circ \mathcal{E}_{1(2)} \cdots \mathcal{E}_{1(2)}(\phi) \\ &= (\mathcal{E}_{1(2)} \otimes I \otimes \cdots \otimes I) \circ (I \otimes \mathcal{E}_{1(2)} \otimes I \cdots \otimes I) \circ (I \otimes \mathcal{E}_{1(2)} \otimes I \cdots \otimes I) \circ \cdots (I \otimes I \otimes \cdots \otimes \mathcal{E}_{1(2)})(\phi) \end{aligned} \tag{14}$$

By induction, the two operations must be perfectly distinguishable in the single-copy case. However this means U_1 and U_2 can be distinguished with one copy that contradicts with the supposition.

Ex 2: If two distinct EB maps $\mathcal{E}_1, \mathcal{E}_2$, $\mathcal{E}_{1(2)}(\rho) = \sum_i |\phi_{i1(2)}\rangle\langle\phi_{i1(2)}| tr(|\psi_{i1(2)}\rangle\langle\psi_{i1(2)}|\rho)$ satisfy $|\langle\phi_{i1}|\phi_{j2}\rangle| \neq 0$ for all i, j, then they are not perfectly distinguishable in any finite-copy case.

Proof.

$$\mathcal{E}_{1(2)}(\rho_{1(2)}) = \sum_i |\phi_{i1(2)}\rangle\langle\phi_{i1(2)}| \otimes tr(|\psi_{i1(2)}\rangle\langle\psi_{i1(2)}|\rho_{1(2)}), \tag{15}$$

If $\mathcal{E}_1(\rho_1)$ is orthogonal to $\mathcal{E}_2(\rho_2)$, then $tr(|\psi_{i1}\rangle\langle\psi_{i1}|\rho_1)$ is orthogonal to $tr(|\psi_{j2}\rangle\langle\psi_{j2}|\rho_2)$. So $tr\rho_1$ is orthogonal to $tr\rho_2$ that means ρ_1 is orthogonal to ρ_2. Contradiction is deduced by similar reasoning as *Ex 1.*

Conjecture: If two EB maps are not perfectly distinguishable in the single-copy case, then they are not in any finite-copy case.

B. Holevo quantity

Analogously, we discuss distinguishability of quantum operations measured by Holevo quantity.

Definition Distinguishability of an ensemble of quantum operations $E = \{\mathcal{E}_i, p_i\}$ is defined as

$$D(E) = \max_{\rho} \chi(\{\mathcal{E}_i \otimes I(\rho), p_i\}). \tag{16}$$

First we simplify the minimization process then show that $D(E)$ is really non-increasing under further quantum operation.

Lemma 2 Given an ensemble of quantum operations $E = \{\mathcal{E}_i, p_i\}$ on d-dimensional Hilbert space, the optimal state to achieve the distinguishability $D(E)$ is a pure state in $d \otimes d$ Hilbert space,

$$D(E) = \max_{|\phi\rangle \in d\otimes d} \chi(\{\mathcal{E}_i \otimes I(\phi), p_i\}). \tag{17}$$

Proof. The proof is similar to that of *Lemma 1.* First a general mixed state ρ^{AB} on $d \otimes n$-dimensional space can always be purified into a pure state $|\Phi\rangle^{ABC}$ with another auxiliary system C, expressed in its Schmidt decomposition splitting between A and BC as,

$$|\Phi\rangle^{ABC} = \sum_{i=0}^{d-1} \sqrt{\lambda_i}|i\rangle^A \otimes |e_i\rangle^{BC}, \tag{18}$$

where $|e_i\rangle^{BC}$ are orthogonal states spanning d-dimensional space. From *Eq.*(2a), we can get

$$\chi(\{\mathcal{E}_i^A \otimes I^B(\rho^{AB}), p_i\}) = \chi(\{tr_C\mathcal{E}_i^A \otimes I^{BC}(\Phi^{ABC}), p_i\}) \leq \chi(\{\mathcal{E}_i^A \otimes I^{BC}(\Phi^{ABC}), p_i\}). \tag{19}$$

Notice that Φ^{ABC} can be transform to the standard form $\sum_{i=0}^{d-1}\sqrt{\lambda_i}|i\rangle^A \otimes |i\rangle^{BC}$ under unitary operator $I^A \otimes U^{BC}$ without varying $\chi(\{\mathcal{E}_i^A \otimes I^{BC}(\Phi^{ABC}), p_i\})$. So the minimization can be considered over pure states in the $d \otimes d$-dimensional Hilbert space. For simple notation, the identity operation is omitted if no confusion appears.

Proposition 2 The distinguishability $D(E)$ is non-increasing under further quantum operations,

$$D(\mathcal{N} \circ E) \leq D(E), \tag{20}$$

where $\mathcal{N} \circ E = \{\mathcal{N} \circ \mathcal{E}_i, p_i\}$ and $\mathcal{N}$ is a quantum operation.

Proof. The proof comes from *Eq.*(2b) and *Lemma 2.* Suppose the optimal pure state to achieve $D(\mathcal{N} \circ E)$ is ϕ^*.

$$D(\mathcal{N} \circ E) = \chi(\{\mathcal{N}(\mathcal{E}_i(\phi^*)), p_i\}) \leq \chi(\{\mathcal{E}_i(\phi^*), p_i\}) \leq D(E) \tag{21}$$

Additivity is a desirable property. In fact, Holevo quantity describes the asymptotic property of $\{\rho_i, p_i\}^{\otimes N}$. Just as Holevo quantity is additive on tensor product of ensembles, it is reduced to the one-copy form. How about the distinguishability defined by *Eq.*(17)? For two ensembles of quantum operations $E_1 = \{\mathcal{E}_i^{Q_1}, p_i\}$ and $E_2 = \{\mathcal{E}_j^{Q_2}, q_j\}$ where E_1 and E_2 operate on d_1-dimensional system Q_1 and d_2-dimensional system Q_2 respectively, the tensor product ensemble is defined as $E_1 \otimes E_2 = \{\mathcal{E}_i^{Q_1} \otimes \mathcal{E}_j^{Q_2}, p_iq_j\}$. Additivity does'nt explicitly hold since for tensor product of two ensembles, the input state is possibly entangled between two ensembles to improve distinguishability of the tensor ensemble. As suggested in [7, 8, 9], an input state entangled with an auxiliary system indeed improves distinguishability between two unitary operators. Also, entangled state can be utilized to better estimate an unknown quantum channel [19, 20]. However, we show that additivity still holds. A direct corollary of *Proposition 2* for two ensembles of operations in sequence is as follows.

Proposition 3 For two ensembles of quantum operations $E = \{\mathcal{E}_i, p_i\}$ and $F = \{\mathcal{F}_i, q_i\}$,

$$D(E \circ F) \leq D(E) + D(F), \tag{22a}$$

$$D(F \circ E) \leq D(E) + D(F), \tag{22b}$$

where $E \circ F = \{\mathcal{E}_i \circ \mathcal{F}_j, p_iq_j\}$ and $F \circ E = \{\mathcal{F}_i \circ \mathcal{E}_j, q_ip_j\}$.

Proof. Suppose the optimal state to achieve $D(E\circ F)$ is $|\Phi^*\rangle$.

$$\begin{aligned} D(E\circ F) &= S(\sum_{ij} p_i q_j \mathcal{E}_i\circ\mathcal{F}_j(\Phi^*)) - \sum_{ij} p_i q_j S(\mathcal{E}_i\circ\mathcal{F}_j(\Phi^*)) \\ &= S(\sum_i p_i \mathcal{E}_i(\sum_j q_j \mathcal{F}_j(\Phi^*))) - \sum_i p_i S(\mathcal{E}_i(\sum_j q_j F_j(\Phi^*))) \\ &+ \sum_i p_i S(\mathcal{E}_i(\sum_j q_j \mathcal{F}_j(\Phi^*))) - \sum_i p_i \sum_j q_j S(\mathcal{E}_i(\mathcal{F}_j(\Phi^*))) \\ &\le D(E) + \sum_i p_i D(\mathcal{E}_i\circ F) \le D(E) + D(F). \end{aligned} \tag{23}$$

The second inequality is proved similarly.

Proposition 4 Distinguishability is additive on tensor product of ensembles of quantum operations

$$D(E_1\otimes E_2) = D(E_1) + D(E_2). \tag{24}$$

Proof. From the definition of D and the additivity of Holevo quantity, it is easy to show $D(E_1\otimes E_2)\ge D(E_1)+D(E_2)$. Notice that $E_1\otimes E_2 = (E_1\otimes I)\circ(I\otimes E_2)$, and $D(E_1\otimes E_2)\le D(E_1)+D(E_2)$ holds by *Proposition 3*.

Now we just know that the Holevo quantity of an ensemble of quantum operations defined in *Eq.*(16) satisfies the constraint of distinguishability, therefore it is a reasonable one. But what's the physical meaning of this quantity? Can it be explained from information-theoretic point in the same way as that of quantum states? The answer is YES. It indeed represents classical information communicated by the ensemble of operations.

C. Classical Information

Before explanation, we briefly review Holevo quantity of an ensemble of quantum states since the reasoning is similar. $\chi(\{\rho_i,p_i\})$ is explained as the classical information that can be conveyed by quantum states as signals [16]. The techniques of block-coding and codeword-pruning are employed in the code-decode process [16]. More precisely, the sender encodes classical messages into long strings of states—codewords; codewords assigned with probability P_s (in fact they are equal) are selected from the the set of strings $\{\rho_{i_1}\otimes\rho_{i_2}\otimes\cdots\rho_{i_n}\}$ satisfying that the codewords are sufficiently distinguishable and the frequency of each letter respects it probability; the receiver decodes the received codeword as a whole. The maximal number of codewords is almost 2^{χ}. This is the information-theoretic explanation of χ. We would like to explain $\chi(E)$ of an ensemble of quantum operations similarly. In this situation, quantum operations act as signal carriers instead of quantum states. The classical messages are encoded into a sequence of black boxes and each box can be inquired only once. We assert that the classical information conveyed by an ensemble of operations is its Holevo quantity.

Proposition 5: Let $\{\mathcal{E}_i,p_i\}$ be an ensemble of quantum operations acting on $d-$dimensional Hilbert space. The achievable classical information conveyed by the ensemble is

$$I(\{\mathcal{E}_i,p_i\}) = \max_{\phi\in d\otimes d}\chi(\{\mathcal{E}_i(\phi),p_i\}). \tag{25}$$

Before the proof, we emphasize that the code-decode process of quantum operations is distinct from that of quantum states though it seems alike. One different point is that a sequence of black boxes can operate in different forms of $\mathcal{E}_{i_1}\otimes\mathcal{E}_{i_2}\circ\mathcal{E}_{i_3}\cdots\mathcal{E}_{i_n}$, where $\otimes$ and $\circ$ can be varied. The other point is that the receiver has the privilege to choose a suitable input state to detect which message the sequence represents. Any entangled state is a choice so that the output state may be entangled among subsystems, i. e. optimization over entangled detector states should be considered. These two facts don't appear in the code-decode process of quantum states. However, complexity is completely solved by *Proposition 3, 4*.

Proof. $I\ge\chi$: The sequence of operations performs as $\mathcal{E}_{i_1}\otimes\cdots\otimes\mathcal{E}_{i_n}$ and the optimal detector state can be chosen as product state as showed by Proposition 4. Then the problem is reduced to that of quantum states by additivity *Eq.*(24).

$I\le\chi$: Suppose there exist a code with N codewords $\mathcal{E}_{i_1}\otimes\mathcal{E}_{i_2}\circ\mathcal{E}_{i_3}\cdots\mathcal{E}_{i_n}$ with probability P_s such that (1) they can be almost distinguishable by ψ and (2) $\mathcal{E}_i$ appears with frequency p_i. Holevo quantity is an upper bound of the classical information of $\{\mathcal{E}_{i_1}\otimes\mathcal{E}_{i_2}\circ\mathcal{E}_{i_3}\cdots\mathcal{E}_{i_n}(\psi),p_{i_1}p_{i_2}\cdots p_{i_n}\}$. In addition to *Proposition 3,4*, we have

$$nI\le\chi(\{\mathcal{E}_i,p_i\}\otimes\{\mathcal{E}_i,p_i\}\circ\{\mathcal{E}_i,p_i\}\cdots\{\mathcal{E}_i,p_i\}(\psi)) \le n\max_{\phi\in d\otimes d}\chi(\{\mathcal{E}_i(\phi),p_i\}). \tag{26}$$

The proof is completed.

Now we show that $\max_{\phi\in d\otimes d}\chi(\{\mathcal{E}_i(\phi),p_i\})$ describes the classical information conveyed by quantum operations. The capacity of classical information conveyed by a set of states $\{\rho_i\}$ is defined as $C(\{\rho_i\})=\max_{p_i}\chi(\{\rho_i,p_i\})$. Analogously, we can define the counterpart of a set of operations.

Definition The capacity of classical information conveyed by a set of operations $\{\mathcal{E}_i\}$ is

$$C(\{\mathcal{E}_i\})=\max_{p_i} I(\{\mathcal{E}_i,p_i\})=\max_{p_i}\max_{\phi\in d\otimes d}\chi(\{\mathcal{E}_i(\phi),p_i\}). \tag{27}$$

Proposition 6 $C(\{\rho_i\}\otimes\{\sigma_i\})=C(\{\rho_i\})+C(\{\sigma_i\})$.

Proof.

$$\begin{aligned}\chi(\{\rho_i\otimes\sigma_j,p_{ij}\}) &= S(\sum_{ij}p_{ij}\rho_i\otimes\sigma_j)-\sum_{ij}S(\rho_i\otimes\sigma_j)\\ &\le S(\sum_i p_{i|\cdot}\rho_i)+S((\sum_j p_{\cdot|j}\sigma_j))-\sum_i p_{i|\cdot}S(\rho_i)-\sum_j p_{\cdot|j}S(\sigma_j),\\ &= \chi(\{\rho_i,p_{i|\cdot}\})+\chi(\{\sigma_j,p_{\cdot|j}\}),\end{aligned} \tag{28}$$

where $p_{i|\cdot}=\sum_j p_{ij}$ and $p_{\cdot|j}=\sum_i p_{ij}$

Proposition 7 $C(\{\mathcal{E}_i\}^{\otimes 2})=2C(\{\mathcal{E}_i\})$.

Proof. Suppose the capacity $C(\{\mathcal{E}_i\}^{\otimes 2})$ is achieved by $\{\mathcal{E}_i\otimes\mathcal{E}_j,p_{ij}\}$ and the optima input ϕ. Then the probability of $\mathcal{E}_i$ is $(p_{i|\cdot}+p_{\cdot|i})/2$. According to the informational explanation of χ, $\chi(\{\mathcal{E}_i\otimes\mathcal{E}_j(\phi),p_{ij}\})$ means that $2^{N\chi}$ codewords can be selected from the set of state strings $\{(\mathcal{E}_{i_1}\otimes\mathcal{E}_{j_1}(\phi))\otimes\cdots\otimes(\mathcal{E}_{i_N}\otimes\mathcal{E}_{j_N}(\phi))\}$ such that the frequency of $\mathcal{E}_i\otimes\mathcal{E}_j(\phi)$ is nearly p_{ij} and the codewords are sufficiently distinguishable—that are measured by the average decoding error. It also amounts to that $2^{2N\cdot\chi/2}$ codewords can be selected from same set with string $2N$ such that the frequency of $\mathcal{E}_i$ occurrence is nearly $(p_{i|\cdot}+p_{\cdot|i})/2$ and the codewords are sufficiently distinguishable. Recall the definition of classical information of ensemble $\{\mathcal{E}_i,(p_{i|\cdot}+p_{\cdot|i})/2\}$, the quantity $\chi/2$ cannot exceed $I(\{\mathcal{E}_i,(p_{i|\cdot}+p_{\cdot|i})/2\})$. So we get $\chi/2\le C(\{\mathcal{E}_i\})$. The other direction is straightforward.

Comparing the classical capacity of quantum operations with one-shot capacity of classical information of a quantum channel $C_1(\mathcal{E})=\max_{\{\phi_i,p_i\}}\chi(\{\mathcal{E}(\phi_i),p_i\})$, we can see that the roles of states and channels are swapped. Also we suffer from the additivity problem.

$$C_1(\mathcal{E}\otimes\mathcal{E}) = 2C_1(\mathcal{E}) \quad ? \tag{29a}$$

$$C(\{\mathcal{E}_i\}\otimes\{\mathcal{F}_i\}) = C(\{\mathcal{E}_i\})+C(\{\mathcal{F}_i\}) \quad ? \tag{29b}$$

$$C_1(\mathcal{E}\otimes\mathcal{F}) = C_1(\mathcal{E})+C_1(\mathcal{F}) \quad ? \tag{29c}$$

D. Unitary Ensemble

Now we focus on the distinguishability of unitary ensemble. From the above discussion, distinguishability of unitary ensemble $\mathcal{U}=\{U_i,p_i\}$ acting on d-dimensional Hilbert space is

$$\begin{aligned}D(\mathcal{U}) &= \max_{\phi\in d\otimes d} S(\mathcal{E}(\phi)),\\ \mathcal{E}(\phi) &= \sum p_i U_i|\phi\rangle\langle\phi|U_i^\dagger.\end{aligned} \tag{30}$$

Note that $\mathcal{E}$ can be regarded as a noisy channel though it is in some special form. Actually, a noisy channel can be expressed as this form if and only if the entanglement of assistance [21] retains invariant after transmission through the channel [22]. In the view of quantum state compression [23], $D(\mathcal{U})$ represents quantum memories required for faithfully storing the output states of the quantum channel. Universal quantum information compression [24] demonstrates that a quantum source that is only known to have von Neumann entropy less than or equal to S but is otherwise completely unspecified, can be faithfully compressed to S qubits by a universal quantum data compression.

First we calculate distinguishability of an ensemble $\{U_1,U_2,p_1,p_2\}$ $U_1,U_2\in SU(d)$. A direct calculation shows that $S(\mathcal{E}(\phi))=S(M)$ where M is a 2×2 matrix $\begin{pmatrix}p_1 & \sqrt{p_1p_2}\langle\phi|U_1^\dagger U_2|\phi\rangle\\ \sqrt{p_1p_2}\langle\phi|U_2^\dagger U_1|\phi\rangle & p_2\end{pmatrix}$. The optimal state to maximize von Neumann entropy is the one to minimize $|\langle\phi|U_1^\dagger U_2|\phi\rangle|$ that amounts to optimally discriminating the two unitary operators. In [8, 9], minimization is achieved as follows. Taking the spectral decomposition of $U=U_1^\dagger U_2$ as $\{|u_i\rangle,u_i\}$,

and $\rho^A = tr_B(|\phi\rangle\langle\phi|)^{AB}$, then $\langle\phi|U_1^\dagger U_2|\phi\rangle = tr(U\otimes I|\phi\rangle\langle\phi|) = tr\sum_i u_i|u_i\rangle\langle u_i|\rho^A = \sum_i \lambda_i u_i$, where $\lambda_i = \langle u_i|\rho^A|u_i\rangle$ is a probability distribution. In the complex plane, a polygon is formed whose vertices is the eigenvalues u_i locating on the circle $|z| = 1$. The minimum $|\sum_i \lambda_i u_i|$ is the shortest distance from the origin to the polygon.

It is not easy to solve the distinguishability for $SU(d)$ ensemble with more than two operators because of two obstacles. One is that the optimal states to discriminate any two operators are not the same in general, and the other is that that even if the same optimal state minimizes all pairwise $|\langle\phi|U_i^\dagger U_j|\phi\rangle|$, it is unnecessarily the optimal one that maximizes von Neumann entropy—Jozsa-Schlienz paradox that occurs almost all ensembles in higher dimensions. However, we will prove the formula of distinguishability of $SU(2)$ ensemble. Fortunately, difficulties are overcome for this simple case. A maximally entangled state is optimal independently of two $SU(2)$ operators to be distinguished [8] and it is also the optimal state to maximize von Neumann entropy although the ensemble of output states is in $2\otimes 2$ space. As well-known, $SU(2)$ is the elementary gate for qubits that is basic system in quantum information theory. $SU(2)$ is studied with detail due to its simplicity and importance. *Proposition 8* is a contribution to $SU(2)$ ensemble.

Proposition 8 Distinguishability of $SU(2)$ ensemble $\mathcal{U} = \{U_i, p_i\}$ is achieved by a maximally entangled state $|\phi^*\rangle = 1/\sqrt{2}(|00\rangle + |11\rangle)$ and the explicit formula is,

$$D(\mathcal{U}) = S(\mathcal{E}(\sum_{i=1}^{n} p_i U_i|\phi\rangle\langle\phi|U_i^\dagger)) = S([\sqrt{p_i p_j}trU_i^\dagger U_j/2]_{n\times n}). \tag{31}$$

Proof. A generic $U \in SU(2)$ matrix can be parameterized by two complex numbers α, β as

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \tag{32}$$

satisfying $|\alpha|^2 + |\beta|^2 = 1$. Suppose the optimal state to achieve distinguishability of $\{U_i, p_i\}$ is $|\psi\rangle = aV|0\rangle\otimes|0\rangle + bV|1\rangle\otimes|1\rangle$ where a, b are nonnegative numbers satisfying $a^2 + b^2 = 1$, and $V \in SU(2)$ is dependent on the ensemble (we can always fix the basis of the auxiliary subsystem in the Schmidt decomposition). Then $|\phi\rangle = a|00\rangle + b|11\rangle$ is the optimal state for $\{U_iV, p_i\}$ and $D(\{U_i, p_i\}) = D(\{U_iV, p_i\})$. Therefore the problem is reduced to the ensemble $\{U_iV, p_i\}$ with optimal state of the form $|\phi\rangle = a|00\rangle + b|11\rangle$. Here V is certainly dependent on $\{U_i, p_i\}$ and still unknown. However, we will show that $|\phi\rangle = 1/\sqrt{2}(|00\rangle + |11\rangle)$ is the universal optimal state for $\{U_iV, p_i\}$ independent of V.

Set $U_iV = W_i \in SU(2)$ and W_i is parameterized by α_i, β_i as the form *Eq.*(32). The ensemble of the output states is $\{|\phi_i\rangle, p_i\}$ where

$$|\phi_i\rangle = W_i|\phi\rangle = aW_i|0\rangle\otimes|0\rangle + bW_i|1\rangle\otimes|1\rangle, \tag{33}$$

and the average state is $\rho = \sum p_i|\phi_i\rangle\langle\phi_i|$. The eigenvalues of ρ are the same as the non-zero eigenvalues of matrix M whose entries are defined as

$$M_{ij} = \sqrt{p_i p_j}\langle\phi_i|\phi_j\rangle. \tag{34}$$

This conclusion [2] can be easily seen by introducing n orthogonal vectors $|e_i\rangle$ in an auxiliary Hilbert space and considering the pure state $|\Upsilon\rangle = \sum\sqrt{p_i}|e_i\rangle|\psi_i\rangle$. ρ and M are just the two reduced states obtained by partial trace of $|\Upsilon\rangle\langle\Upsilon|$ over the first and second components respectively. It follows that the Hermite matrix M has the same eigenvalues as ρ .

As $W_i^\dagger W_j \in SU(2)$, the two diagonal elements of $W_i^\dagger W_j$ are conjugate numbers denoted as α_{ij} and α_{ij}^*. M_{ij} is explicitly written as

$$M_{ij} = \sqrt{p_i p_j}(a^2\langle 0|W_i^\dagger W_j|0\rangle + b^2\langle 1|W_i^\dagger W_j|1\rangle) = \sqrt{p_i p_j}(a^2\alpha_{ij} + b^2\alpha_{ij}^*). \tag{35}$$

We demonstrate that maximum of $S(M)$ is achieved when $a^2 = b^2 = 1/2$ is satisfied, i. e. $|\phi^*\rangle$ is a maximally entangled state. This is concluded from the following reason. If the input state is maximally entangled state $|\phi^*\rangle$, its corresponding matrix G is

$$G_{ij} = \frac{1}{2}\sqrt{p_i p_j}(\alpha_{ij} + \alpha_{ij}^*) = \frac{1}{2}M_{ij} + \frac{1}{2}M_{ij}^*, \tag{36}$$

that means $G = \frac{1}{2}M + \frac{1}{2}M^*$. Notice that M^* is also a Hermite matrix with the same eigenvalues as M. Therefore there exists a unitary transformation T satisfying $M^* = TMT^\dagger$. Now G can be written as,

$$G = \frac{1}{2}M + \frac{1}{2}TMT^\dagger. \tag{37}$$

It immediately follows from Uhlmann theorem that $\lambda(G) \prec \lambda(M)$, which means that the eigenvalues of G is majorized by those of M [25, 26]. As a result of the theory of majorization, $S(G) \geq S(M)$.

Write G_{ij} explicitly,

$$G_{ij} = \frac{1}{2}\sqrt{p_i p_j}(\langle 0|W_i^\dagger W_j|0\rangle + \langle 1|W_i^\dagger W_j|1\rangle) = \frac{1}{2}\sqrt{p_i p_j} tr W_i^\dagger W_j = \frac{1}{2}\sqrt{p_i p_j} tr(U_i V)^\dagger U_j V = \frac{1}{2}\sqrt{p_i p_j} tr U_i^\dagger U_j. \quad (38)$$

It is clear that G is independent of V. So $|\phi^*\rangle$ is the universal optimal state to achieve distinguishability of $SU(2)$ ensemble and the formula is obtained.

Just as Jozsa-Schlienz paradox that appears in an ensemble of quantum states [2], we show it also occurs for an ensemble of quantum operations for the minimal dimensional case—for $SU(2)$ ensembles. It is possible to increase distinguishability of all pairwise but to decrease the global distinguishability. Notice that the phenomenon cannot occur for ensembles of quantum states in 2-dimensional space. The paradox in case of $SU(2)$ ensemble is illustrated as follows.

Ex 3 Consider two $SU(2)$ ensembles, $\mathcal{U} = \{U_1, U_2, U_3, 1/3, 1/3, 1/3\}$ and $\mathcal{V} = \{V_1, V_2, V_3, 1/3, 1/3, 1/3\}$ where

$$U_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, U_2 = \begin{pmatrix} \sqrt{1/2} & \sqrt{1/2} \\ -\sqrt{1/2} & \sqrt{1/2} \end{pmatrix}, U_3 = \begin{pmatrix} \sqrt{1/3} & \sqrt{2/3}e^{-i\alpha} \\ -\sqrt{2/3}e^{i\alpha} & \sqrt{1/3} \end{pmatrix}$$

$$V_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, V_2 = \begin{pmatrix} \sqrt{1/2.1} & \sqrt{1.1/2.1} \\ -\sqrt{1.1/2.1} & \sqrt{1/2.1} \end{pmatrix}, V_3 = \begin{pmatrix} \sqrt{1/3.1} & \sqrt{2.1/3.1}e^{-i\beta} \\ -\sqrt{2.1/3.1}e^{i\beta} & \sqrt{1/3.1} \end{pmatrix}$$

in which $cos(\alpha) = \frac{\sqrt{3}}{4} - \frac{1}{\sqrt{2}}$ and $cos(\beta) = -\frac{1}{\sqrt{2.1 \times 1.1}}$.

Now it is easy to verify

$$\begin{aligned} |tr U_1^\dagger U_2/2| = 1/\sqrt{2} &> |tr V_1^\dagger V_2/2| = 1/\sqrt{2.1}, \\ |tr U_1^\dagger U_3/2| = 1/\sqrt{3} &> |tr V_1^\dagger V_3/2| = 1/\sqrt{3.1}, \\ |tr U_2^\dagger U_3/2| = 1/4 &> |tr V_2^\dagger V_3/2| = 0, \\ D(\mathcal{U}) \approx 1.138 &> D(\mathcal{V}) \approx 1.118, \end{aligned}$$

which means that each pair of $\mathcal{U}$ is less distinguishable than the corresponding pair of $\mathcal{V}$, yet as a whole $\mathcal{U}$ is more distinguishable than $\mathcal{V}$.

V. CONCLUDING REMARKS

In summary, we show how to define two self-consistent measures of distinguishability under a basic property of distinguishability. One is fidelity from the intuition that the more alike the less distinguishable, and the other is Holevo quantity from the information-theoretic consideration. First, distinguishability of pure states is extended to mixed ones. Then based on distinguishability of quantum states, distinguishability of quantum operations is defined as the largest distinguishability of output states. Especially, Holevo quantity of an ensemble of quantum operations is explained from the information-theoretic point. Properties of Holevo quantity of operations are discussed. The analytic formula for computing distinguishability of $SU(2)$ ensemble is proved. With the aid of the formula, we show that Jozsa-Schlienz paradox also appears in quantum operations.

[1] R. Jozsa, J. Mod. opt. **41**, 2315 (1994).
[2] R. Jozsa and J. Schlienz, Phys. Rev. A **62**, 012301 (2000).
[3] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[4] M. A. Nielsen, C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow, and A. Hines, Phys. Rev. A **67**, 052301 (2003).
[5] M. D. Choi, Linear Algebra Appl. **10**, 285 (1975).
[6] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[7] A. M. Childs, J. Preskill and J. Renes, J. Mod. Opt. **47**, 155 (2000).
[8] A. Acin, Phys. Rev. Lett. **87**, 177901 (2001)
[9] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).
[10] A. S. Kholevo, Probl. Peredachi Inf. **9**, 110 (1973) [Proble. Inf. Transm. (USSR) **9**, 31 (1973)].

[11] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[12] G. Lindblad, Comm. Math. Phys. **40**, 147 (1975).
[13] A. Uhlmann, Comm. Math. Phys. **54**, 21 (1977).
[14] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Commun. Math. Phys. **246(2)**, 359 (2004).
[15] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[16] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[17] Michael Horodecki, Peter W. Shor, Mary Beth Ruskai, Rev. Math. Phys. **15**, 629 (2003).
[18] G. M. D'Ariano and P. Lo Presti, Phys. Rev. Lett. **86**, 4195 (2001).
[19] A. Acin, E. Jane, and G. Vidal, Phys. Rev. A **64**, 050302 (2001).
[20] G. M. D'Ariano and P. Lo Presti, Phys. Rev. Lett. **91**, 047902 (2003).
[21] D. P. DiVincenzo, C. A. Fuchs, H. Mabuchi, J. A. Smolin, A. Thapliyal, A. Uhlmann, *Entanglement of Assistance*, quant-ph/9803033.
[22] D. Yang, in preparation.
[23] B. Schumacher, Phys. Rev. A **51**, 2738 (1995); R. Jozsa and B. Schumacher, J. Mod. Opt. **41**, 2343 (1994).
[24] R. Jozsa, M. Horodecki, P. Horodecki, R. Horodecki, Phys. Rev. Lett. **81**, 1714 (1998).
[25] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[26] R. Bhatia, *Matrix Analysis* (Springer-verlag, New York, 1997).